2018

Evaluation of Consumer Drone Control Interface

Thomas William Merrell
Wright State University

EVALUATION OF CONSUMER DRONE CONTROL INTERFACE

A thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science in Industrial and Human Factors Engineering

By

THOMAS W. MERRELL, JR.
B.S.B.E., Wright State University, 2014

2018
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

December 14, 2017

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Thomas W. Merrell, Jr. ENTITLED Evaluation of Consumer Drone Control Interface BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Industrial and Human Factors Engineering.

Subhashini Ganapathy, Ph.D.
Thesis Director

Jaime E. Ramirez-Vick, Ph.D.
Chair, Department of Biomedical, Industrial and Human Factors Engineering

Committee on
Final Examination

Subhashini Ganapathy, Ph.D.

Mary Fendley, Ph.D.

Sasanka Prabhala, Ph.D.

Barry Milligan, Ph.D.
Interim Dean of the Graduate School

Abstract

Merrell, Thomas W., Jr. M.S.I.H.E. Department of Biomedical, Industrial, and Human Factors Engineering, Wright State University, 2018. Evaluation of Consumer Drone Control Interface.

The development and use of consumer grade drones is becoming a larger part of our society for many different applications. There has been a great amount of discussion and constant review of proper operation of consumer drones including proper methods of control. In turn, regulation of such devices has been inconsistent. This study aims to better understand the effects of the three primary control interface methods (line of sight, video aided, and first-person view) on flight performance, situational awareness, and perceived mental workload of the operator. Secondarily, this study aims to provide design recommendations for future interfaces. This study shows that the first-person view control interface results in a longer flight time around a course, higher mental workload, and lower situational awareness when compared to line-of-sight and video aided control. The use of line-of-sight control performed superiorly in all areas, and the video-aided interface was very close behind.

Table of Contents

List of Figures

List of Tables

1. BACKGROUND

This section reviews existing literature on the challenges of operating a consumer grade quadcopter (drone), operator mental workload, situational awareness, and the Federal Aviation Administration rules, regulations, and guidelines.

1.1 Consumer Unmanned Aerial Vehicles

Consumer unmanned aerial vehicles, otherwise known as drones, are becoming more and more popular. The number of consumer drone shipments have risen from approximately 3 million in 2014 to 7 million in 2016 and are projected to reach 29 million by 2021 (Meola, 2017). The Business Insider has defined a consumer drones as “aerial vehicles that can fly autonomously or be piloted by remote individual” (Meola, 2017). This only includes drones purchased for personal/non-professional use and not those purchased for professional/commercial use. This is indeed a very large and profitable market with sales of consumer drones in 2017 at approximately $1.3 billion dollars in the United States alone (Statista, 2017).

Many small consumer drones contain limited available onboard sensory equipment, which can lead to low altitude operation near populated areas and this can result in unforeseen interactions between people and drones (Magister, 2010). The risk of drone-human collisions is of increasing interest. These interactions have been modeled using blunt ballistic impact and there is evidence that these interactions could become lethal if a human is struck by a sharp part of the drone (Magister, 2010). Reports of severe injuries involving novice operators losing control and colliding with people have been documented. In one case, a drone struck a child and caused serious injuries to the face and eye, resulting in multiple surgeries (BBC News, 2015). The risks are very real and to better prevent accidents such as this one it is necessary to understand how the various control interface methods effect the operator with respect to mental workload and situational awareness. Currently, there are 3 common types of control interfaces that are available with consumer drones which include line-of-sight (LOS), video aided, and first-person-view (FPV). LOS is the most commonly used control interface for consumer drones. Every drone has the capability to be operated via LOS. LOS control requires that the operator can physically see the drone in order to operate it. This method is also recommended by the Federal Aviation Administration over other methods. For an example of LOS control, see figure 1.

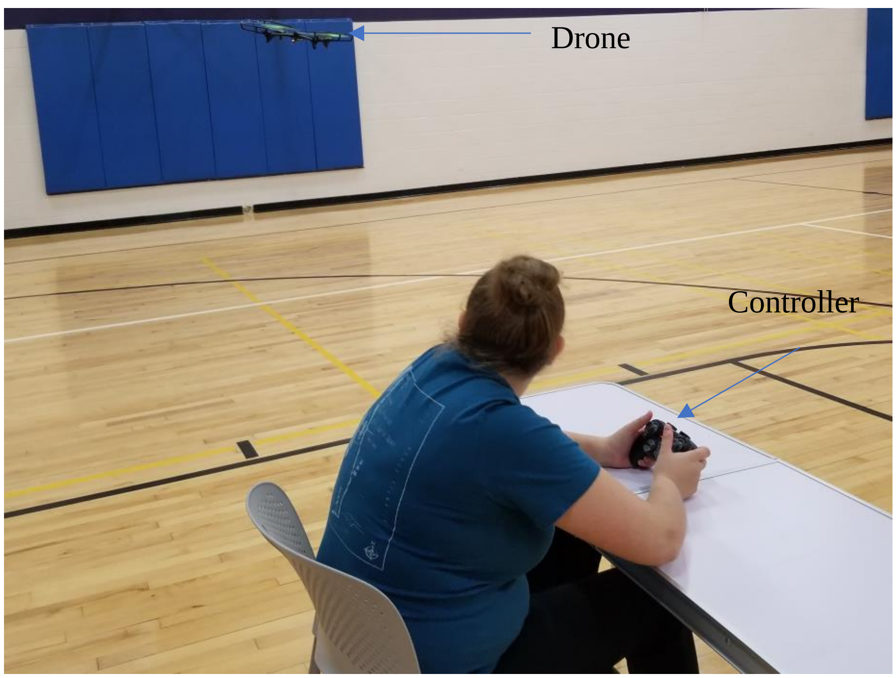

Figure 1: Line-of-sight control interface method. The operator manipulates the controller while keeping their eyes on the drone.

Many drones are now being outfitted with forward facing cameras, which can be used to stream video to a mobile device or computer. These cameras are key components of video-aided and first-person-view control interface methods. Video-aided control is a combination of LOS control and video streaming in which the operator can maintain the drone in physical LOS while streaming video from a forward-facing camera on the drone. This video is typically streamed to a mobile device such as a mobile phone or tablet. In figure 2, an example of video-aided control can be seen where the operator is streaming the video from the drone to a monitor. This kind of video streaming allows the operator to utilize their own field of vision as well as that of the drone camera to navigate through the environment.

Figure 2: Video aided control interface method. The operator can see the video stream from the front mounted camera on the drone, as well as maintain visual line of sight.

With the inclusion of the forward-facing cameras on drones, the use of FPV control is also growing. This control interface method streams video from the drone to a headset that the operator wears. First-person view is the official method of control used in the Drone

Racing League. This method allows the operator to solely use the perspective of the drone's front mounted camera, while maneuvering through the environment. Figure 3 below shows a participant operating the drone with the FPV control interface.

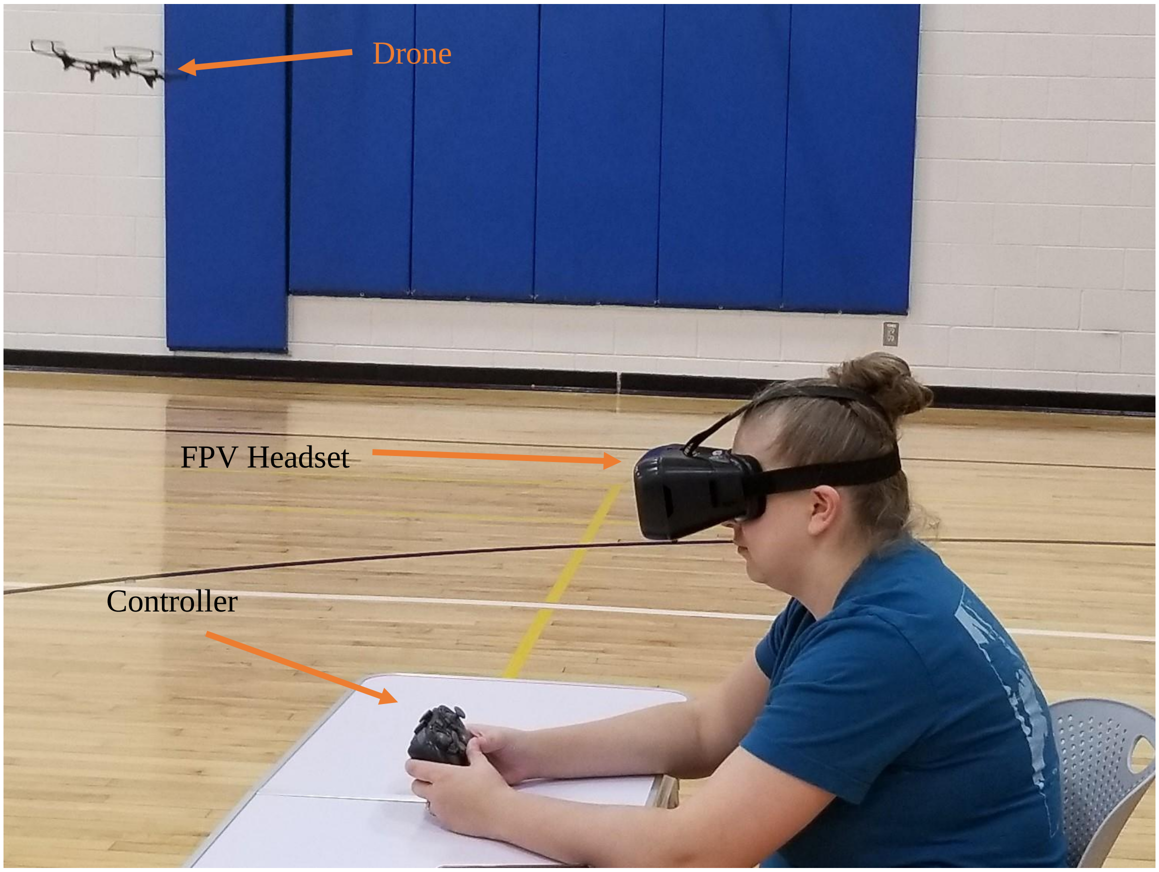

Figure 3: First-person-view control interface method. The operator is wearing an FPV headset which streams the video shown in the video aided interface and replaces the operators own field of view with that of the drone.

1.2 Federal Aviation Administration Rules, Regulations, and Guidelines

The operation of unmanned aircrafts, or drones, falls under the authority of the Federal Aviation Administration (FAA). There has been a great deal of controversy involving the operation of drones for both consumer and commercial use and their interference with other aircraft.

From January 2016 to December 2016, there were over 1000 incidents involving drones interacting with other aircrafts, which were reported to various law enforcement agencies and the FAA across the United States (Federal Aviation Administration, 2017). The FAA does not confirm that these sightings are actually drones, though they are perceived to be by the public and professionals operating the aircrafts. In addition, most operators do not have a pilot's license and are operating unregistered drones. To reduce the number of these interactions, the FAA has released rules and regulations regarding the operation of consumer drones which are currently under examination to determine if the FAA has the legal power to institute and enforce such standards. A recent appeal has determined that the measures the FAA has taken to register and monitor consumer drones is unlawful. It has now been decided that consumers do not have to register their drones or display identification numbers on them in order to operate them as long as they maintain visual line of sight control. This means that the operator or a co-pilot must maintain the visual contact with the drone (John A. Taylor V. Michael P. Huerta, 2017). Though there is still contention as to the lawfulness of the FAA restricting flight space for drones and other model aircrafts.

1.3 Challenges in Human Factors Engineering

Some of the challenges that are presented when designing a control interface for a consumer drone, or determining which available interface to use, revolve around human perception and cognition. One such challenge is the perceived mental workload of the task, which can be considered as task difficulty. Another challenge is situational awareness, which is how well the operator understands their current situation. Flight accidents are frequently found to be related to situational awareness (Lu, Horng, & Chao, 2013). These topics are discussed in further detail in the following sections.

1.4 Mental Workload

Workload is a concept that represents the cost of meeting the requirements for a task. This concept is particularly useful, since there is no way for an operator to perform a task with perfect accuracy every time, there is a need to evaluate how the operator is performing across stages of the task (Hart, 2006). Mental workload is a subjective experience, which is based the task requirements, as well as the circumstances surrounding the task and the operator's skills, behaviors, and perceptions (Hart & Staveland, 1988).

This is a key component in the development and design of any control interface. If the interface increases the level of mental workload for a given task, then there is no benefit to using that interface since it only makes the task more difficult for the operator. In terms of interface design for consumer drones, the designer must consider the available sensors and systems for the drone so that the relevant available information can be displayed to the operator. There are many studies that evaluate alternative interface designs for drones, but little has been done to examine the cognitive and perceptual effects of the three most common interface types. For example, LaFleur et al. (2013) examined the use of electroencephalogram to develop a brain-computer interface with which the operator can control the movement of the drone. (LaFleur, et al., 2013). Cho et al found that an egocentric control interface, which focuses on the operator's perspective rather than a drone-centric interface (video from drone), increases performance (Cho, Cho, & Jeon, 2017). This focuses on the egocentric interface (operator oriented) when compared to the perspective of the video from the drone only and not the line-of-sight control or the video aided control which provides the operator with the drone-centric view as well as the operator's perspective from line-of-sight. There are also several papers, such as Lu and Lung (2016), which focus on the incorporation of gestural control of drones, (Lu & Lung, 2016). These studies find that the use of a device, such as a Microsoft Kinect, to capture the movement of an operator and convert the motion of the operator to actions for the drone, are focused on the operation of the drone, not on the flight performance.

Hooey et al. developed a taxonomy to classify the drivers of mental workload in unmanned vehicle systems. Drivers were classified as environment, task, equipment, or operator (Hooey, Kaber, Adams, Fong, & Gore, 2017). The environmental factors that affect the operation of a drone include both the environment that the drone is operating in and the control environment (Hooey, Kaber, Adams, Fong, & Gore, 2017). This means that environmental conditions such as weather effect the operator's mental workload as well as the environment in which the operator is working from. The next class in this taxonomy is the task itself. The task class contains three subclasses of driver, which include task demands, temporal demands, and task structure (Hooey, Kaber, Adams, Fong, & Gore, 2017). Task demands considers how critical the task is as well as how severe the consequences are. Temporal demands considers how quickly events will arise during the task and task structure considers many different aspects of a single or multi-task event the operator must monitor to successfully accomplish the task(s) (Hooey, Kaber, Adams, Fong, & Gore, 2017). The next class of driver is the equipment. This includes the type of drone, the payload, the onboard sensors, as well as the command, control, and communication link (Hooey, Kaber, Adams, Fong, & Gore, 2017). This focuses on how the operator receives information from, and communicates commands to the drone, and how the drone will respond to those commands. Finally, the operator is the final major class of workload driver in the taxonomy presented by Hooey et al. (2017). This class focuses on how skilled the operator is and the individual differences between the operators. This class cannot be controlled for in the design of consumer drone control interfaces thus the interface should be designed to accommodate the operator, rather than assuming that the operator will be proficient with a particular method of control. Considering mental workload and commonly used interface types could optimize the design of new control interfaces.

1.5 Situational Awareness

Situational awareness (SA) has a key role in the human operator's performance during operation of a drone, and poses major challenges to human performance since human cognition is selective and limited (Raymond, Shapiro, & Arnell, 1992). For operators to quickly recognize that a problem has arisen, they must maintain a high level of situational awareness (Tharanathan, Bullemer, Laberge, Reising, & Mclain, 2012). Situational awareness has been defined as the perception of elements in the environment, comprehension of their meaning, and, at least in the short-term, projection of their future status (Kaber, Jin, Zahabi, & Pankok, Jr., 2016) (Endsley M. R., 1995). Situational awareness can be considered at three levels (Endsley M. R., 1995):

- Level 1: Perception – The operator can perceive the process conditions.
- Level 2: Comprehension – The operator can integrate the perceived information to understand the current state of the process.

- Level 3: Projection – The operator can foresee what the status of the process will be in the next several minutes, including the results of an intervention.

The above levels of SA are not a chronological progression from level 1 to level 3, and instead are considered ascending levels of increased SA (Satuf, Kaszkurewicz, Schiru, & de Campos, 2016). This means that one can perceive the environment without understanding what it means (Lu, Horng, & Chao, 2013) (Satuf, Kaszkurewicz, Schiru, & de Campos, 2016). It has been established that there is a relationship between situational awareness and working memory, time-sharing ability, and perceptual skill (Kaber, Jin, Zahabi, & Pankok, Jr., 2016).

Situational awareness can be thought of in two parts: the process and the product. The process of situational awareness can be thought of as the cognitive processes that lead to the comprehension of elements in the environment.. The product of awareness is the retention of information that can be passed on or assessed (Durso & Sethumadhaven, 2008). Such information retention, however, does not assume that an individual comprehends the information that has been presented to them (Durso & Sethumadhaven, 2008). One phenomenon that can have a significant negative effect on an operator's flight performance is change blindness. This occurs when events in the environment are unexpected or occur outside the focus of attention (Raymond, Shapiro, & Arnell, 1992). Change blindness might arise when visual stimuli are not sufficiently salient enough to be detected. Change blindness can also occur when the visual stimuli are sufficiently salient (Boring, Ulrich, & Lew, 2016). The operator's focus may be on a different area of the control task when the relevant changes to the environment occur (Boring, Ulrich, & Lew, 2016). This has been shown to be detrimental to operator performance in supervisory control tasks, leading operators to miss unexpected changes in the environment, especially when there is another event occurring at the same time (Parasuraman, Cosenzo, & de Visser, 2009). The interruption of tasks can also lead to change blindness by causing the operator to lose focus on the primary goal and switch to a different task. This is a point that interface designers must take into account, as it presents a major potential cause of human error.

2. RESEARCH OBJECTIVES

2.1 Research Objectives

The purpose of this study is to better understand the effects of three visual control interfaces used for the control of a consumer drone on the situational awareness (SA), perceived mental workload, flight performance (lap time, number of collisions, and number of course deviations), as well as the usability of each of the control interfaces. The effects of the type of throttle used (manual vs. automatic) and the number of obstacles (two vs. four) on the course on the flight performance, SA, and perceived mental workload are considered as well. The following are the hypotheses for this study:

1. The use of the first-person view control interface will result in significantly better flight performance than line-of-sight and video aided control interfaces.

    H_0: There is no significant difference in flight performance while using the different control interface types.

    H_1: Operators will experience significantly better flight performance while using the first- person view control interface compared to line-of-sight and video aided control interfaces.

2. The use of the first-person view control interface will result in a significantly higher level of situational awareness than the line-of-sight control and video aided control.

    H_0: There is no significant difference in situational awareness while using the different control interface types.

H_1: Operators perceive a higher level of situational awareness while using the first-person view control interface, compared to the line-of-sight and video aided control.

3. The use of the first-person view control interface will result in significantly lower perceived mental workload than when using the line-of-sight or video aided control.

    H_0: There is no significant difference in mental workload while using the different control interface types.

    H_1: Operators perceive a significantly lower mental workload while using the first-person view control interface, compared to the line-of-sight interface and the video aided control interface.

4. The first-person view control interface is significantly more usable than the line-of-sight or video aided control.

    H_0: There is no significant difference in usability of the different control interface types.

    H_1: The first-person view control interface is significantly more usable than the line of sight control and video aided control.

3. METHODOLOGY

3.1 Experimental Design

The experiment was designed to determine the effects of three control interface methods on the SA, mental workload, and flight performance of users while navigating a track with complexity due to obstacles. The independent variables included the control interface, throttle type and course complexity, while the dependent variables evaluated included the flight performance (lap time, number of course deviations, and number of collisions), SA, perceived mental workload, and usability. The SA, mental workload, and usability were evaluated using surveys, while the flight performance was evaluated via observation.

3.2 Independent Variables

3.2.1 Control Interface Type

The control interface was divided into three types: line of sight, video and line of sight, and first-person view, see figure 4. Line of sight required the operator to only rely on what could be seen from their position at the starting line to discern drone orientation and position on the course. These control interfaces were chosen because they are the three standard interface designs for a consumer drone, with first-person view also used in professional drone racing.

Figure 4: A) LOS control interface method. B) Video aided control interface method. C) FPV control interface method. D) A generic view for the operator while using FPV to operate the drone.

3.2.2 Throttle Type

The type of throttling that the operator used was divided into two categories: automatic and manual. Manual throttle required the operator to monitor and adjust the power of the rotors to reach and/or sustain elevation, while automatic throttle used the on-board controller to maintain an average elevation regardless of the position of the throttle stick, the automatic throttle elevation could be raised or lowered by pressing a button. These two methods come standard on all consumer drones, either together or separate.

3.2.3 Course Type

The course (shown in figure 5 below) that the operators had to navigate was also divided into two categories: simple and complex. Both courses were the same length and width, but the simple course had only two obstacles for the operator to avoid on the track and the complex course had four obstacles on the track for the operator to avoid.

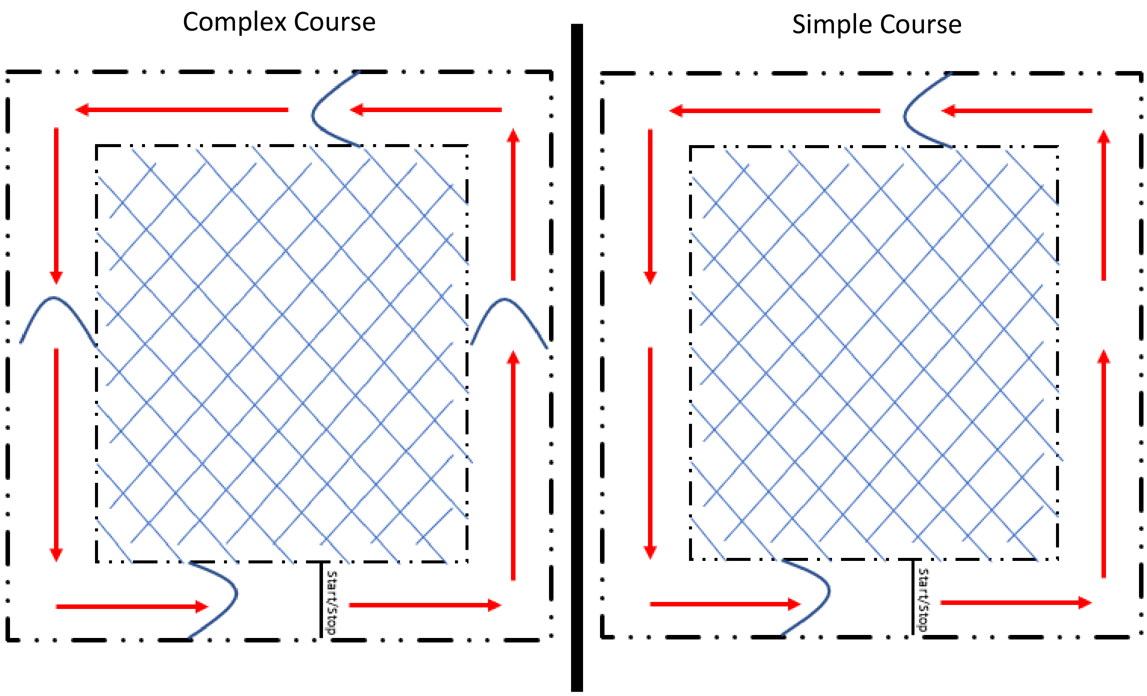

Figure 5: The complex course layout is shown on the left and the simple course is shown on the right. The simple course consists of only two obstacles, while the complex course consists of four obstacles.

3.3 Dependent Variables

3.3.1 Flight Performance

The flight performance is considered in three parts: the lap time, number of collisions, and number of course deviations. Lap time, number of collisions, and number of course deviations are recorded via observation. The lap time is the result of how long the operator takes to maneuver around the course from the start line to the stop line. The number of collisions considers anytime the operator maneuvers the drone into an obstacle on the course or makes contact with an object in the surrounding environment, such as the walls of the gym or the course markers. The number of course deviations considers when the operator maneuvers the drone off of the designated course, outside of the course markers. A stop watch was used to determine the lap time, while simply observing the lap was used to determine the number of collisions and course deviations.

3.3.2 Surveys

There are three surveys that will be used to determine the perceived situational awareness, mental workload, as well as the usability of the different control interface types. All three of the surveys were scaled to have scores of 0 (low) to 100 (high).

3.3.2.1 NASA-TLX

Subjective measures, such as NASA-TLX are very important tools used for the evaluation of systems, and are used extensively for the assessment of mental workload (Rubio, Diaz, Martin, & Puente, 2004). According to Rubio et al. the suitability of an evaluation method for mental workload depends on the following criteria:

1. Sensitivity – The tool’s ability to detect changes in task difficulty/demand.
2. Diagnosability – The identification of changes in workload, as well as the reason for the changes.
3. Selectivity/Validity – The tool should be sensitive only to mental workload.
4. Intrusiveness – The tool should not interfere with the primary task.
5. Implementation Requirements – What is needed to implement the tool.
6. Reliability – The tool should consistently reflect the mental workload.

7. Subject Acceptability – How useful does the subject perceive the tool to be.

NASA-TLX evaluates the subject's perceived mental workload across six dimensions: mental demand, physical demand, frustration, temporal demand, performance, and effort (NASA, 2017) (Rubio, Diaz, Martin, & Puente, 2004). The traditional NASA-TLX is scored by subjective pairwise comparisons of the six dimensions which weighs the individual dimensions, recently there has been increased use of the raw scores, with no weighting, which has been shown to be at least as effective as the traditional survey, and is possibly more indicative of the mental workload (Hart, 2006).

3.3.2.2 Situational Awareness Rating Technique

The situational awareness rating technique (SART) is an assessment of the operator's situational awareness based on the operator's subjective opinion (Endsley et al, 1998). SA is broken up into three components in the SART survey (1) demand, (2) supply, and (3) understanding (Endsley et al, 1998). One of the primary advantages of the SART survey is that it can be administered easily with little to no modification, while a drawback is that it is subjective and thus leaves it up to the operator to account for what they don't know about a situation (Endsley et al, 1998).

3.3.2.3 System Usability Score

The system usability scale (SUS) was first developed in 1986 and remains one of the most widely used and reliable ways to determine usability, and learnability (Brooke, John 1996; U.S. Department of Health and Human Services 2017). SUS is an industry standard, it can be used with small sample sized with reliable results. It is also able to effectively differentiate between usable and unusable systems (Brooke, John 1996; U.S. Department of Health and Human Services 2017). The scoring is out of 100 with the average product scoring a 68, anything higher than 68 is above average and anything below 68 is below average, and should be considered by normalizing the scores to produce a percentile ranking (U.S. Department of Health and Human Services 2017).

3.4 Recruitment

Wright State University undergraduate and graduate students who have normal or corrected to normal vision, and experience using a gaming style controller. Criteria for exclusion included cognitive impairment, and physical impairment that will impede the use of the flight controller. A total of 20 subjects were recruited (10 male, 10 female). Subjects were taken to the WSU Student Union Gymnasium, where a track was prepared. Participants were asked to sign the informed consent document and then shown the three types of surveys that will be administered throughout the test. Each survey was explained to the participants so that they understood how to mark and what the questions meant. They were also told that if they had any questions about the survey at any time to ask.

3.5 Testing Procedure

Each participant was then given instruction on how to operate a SkyViper v2400fpv drone (shown in figure 6) by the experimenter, with the operation demonstrated. Once the controls and operation had been demonstrated to the participant, he/she was able to practice operation for one hour. Participants could freely switch between the three control types, as well as the automatic or manual throttle methods. They were not able to fly the track prior to testing to reduce learning effects. The experimenter provided advice and answered any questions participants had during practice.

After the one-hour practice period participants were asked if they were ready to begin. The trials were balanced using a randomized full factorial design so that each participant experienced all twelve combinations of control interface type, course complexity, and throttle type in a balanced randomized order. Once ready the drone was lined up at the starting line. For each lap, the experimenter gave the start signal and the

Figure 6: The SkyViper v2400fpv drone that is used in this study, as well as the controller and the FPV headset.

drone took-off once it passed the start line the time began, the time stopped when the drone crossed the start line again. The number of course deviations (when the operator leaves the marked course) and collisions (when the operator collides with the course markers, obstacles and the surrounding environment) was recorded by observation. Immediately following a lap, the participant was given a situational awareness rating technique survey and a NASA-TLX survey. At the completion of all twelve laps three system usability scale surveys were administered, one for each of the control interface methods (LOS, video aided, and FPV).

3.6 Statistics

The significance of the factors of control interface type, throttle type, and course type are determined using analysis of variance (ANOVA) with the statistical analysis software JMP from the SAS Institute Inc. A significance of ($\alpha = 0.05$). Along with the ANOVA the interactions of the factors were also examined to determine if there was a significant effect from the factors on the responses.

Each of the three surveys (NASA-TLX, SART, and SUS) were rated on a scale of 0 to 100, with 0 being low and 100 being high. The scores for the SUS survey can be considered in three ranges: below average (< 67), average ($= 68$), and above average (> 68).

4. RESULTS

4.1 Flight Performance

The flight performance results indicate that the type of control interface significantly impacts the operator's lap time ($F = 4.6514$, p-value $= 0.0106$), the mean lap times for each control interface across each throttle type and course type are shown below in table 1. The average lap time when using the FPV interface (57.96 seconds) was significantly higher than the video aided and the LOS (46.18 and 44.47 seconds respectively).

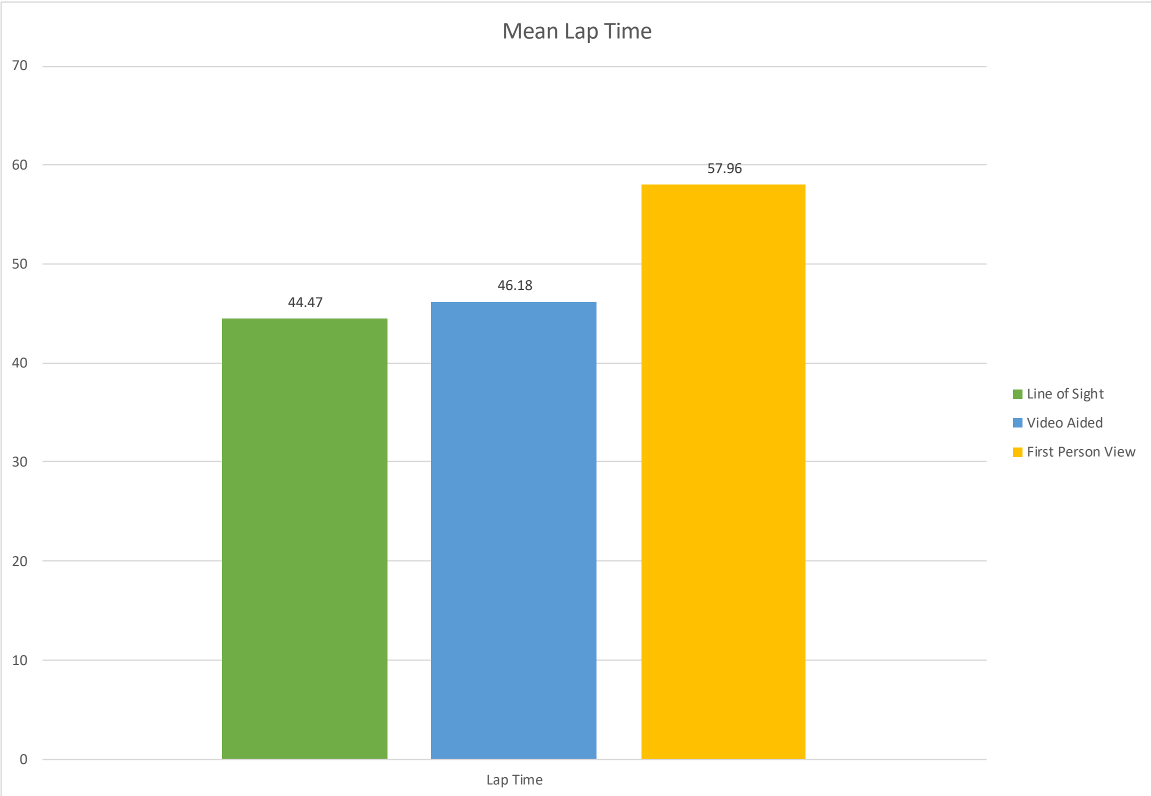

Figure 7: The average lap time for each of the control interfaces.

The results also indicate that there is not a significant difference across control interface types and the number of collisions (F = 2.4053, p-value = 0.0926). The number

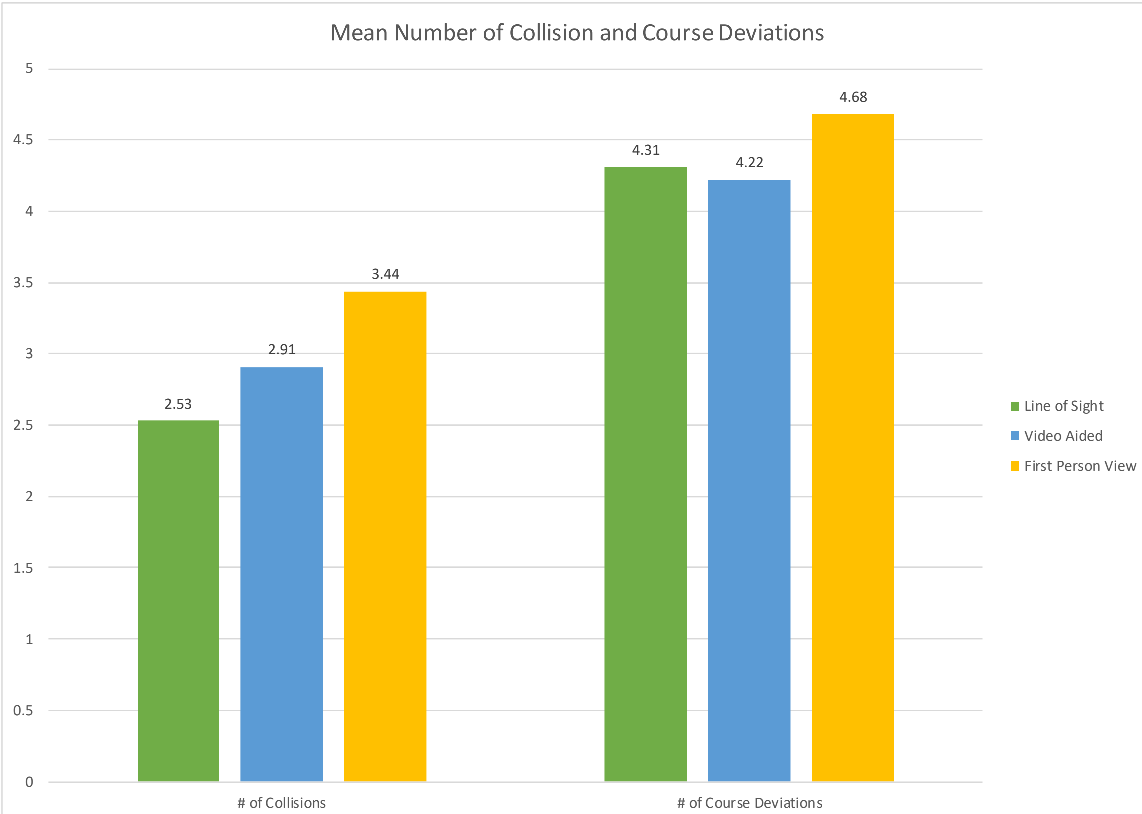

Figure 8: Average number of collisions and course deviations for each interface type.

of course deviations is the final component of flight performance, and the results indicate that there is no significant difference in the number of course deviations when using the different control interface types (F = 1.6464, p-value = 0.1952) and the mean number of course deviations is shown below in table 1.

Further analysis was conducted examining the effects of the throttle control type and the number of obstacles on the flight performance. All results show that there is no significant effect from the number of obstacles on any of the flight performance metrics. The throttle control type, however, did show a significant effect on the lap time (F = 23.6618, p-value < 0.0001), and the number of collisions (F = 25.8414, p-value = <

0.0001), and no significant effect on the number of course deviations (F = 3.3765, p-value = 0.0676).

Table 1: Summary of the data for each of the twelve combinations of control interface, throttle type, and course type.

Control Interface Type	Throttle Type	Course Type	Lap Time (sec/lap)	Number of Collisions (per lap)	Number of Deviations (per lap)
LOS	Auto	Simple	31.44	1.5	2.5
		Complex	37.85	1.32	3.79
	Manual	Simple	47.43	2.58	4.37
		Complex	58.02	4	5.21
Video Aided	Auto	Simple	41.42	2.33	3.94
		Complex	42.48	2.58	4.84
	Manual	Simple	49.24	3.15	4.1
		Complex	51.19	3.53	4
FPV	Auto	Simple	42.94	2.17	4.17
		Complex	42.72	2.11	4.53
	Manual	Simple	79.09	4.89	5.32
		Complex	66.29	4.53	4.68

There is also no significant effect from the interaction between the control type and throttle type for the lap time (F = 2.8059, p-value = 0.0627), number of collisions (F = 2.5836, p-value = 0.0779) and the number of course deviations (F = 2.8282, p-value = 0.0614).

4.2 Perceived Mental Workload

The operators' post lap NASA task load index surveys on perceived mental workload showed a significantly higher mental workload when the operator was using the FPV control interface compared to the video aided and the LOS control interfaces ($F = 6.3903$, $p\text{-value} = 0.0020$). The means (table 2) were examined further with a Tukey-Kramer pairwise comparison which showed that there was a significantly higher mental workload while using FPV than there was while using LOS, and the video aided was not significantly different from either of the other interfaces.

There was no significant effect from the number of obstacles on the perceived mental workload of the operators. The mean perceived mental workload with only two obstacles was 52.76 and the mean with four obstacles was 54.09, both of which are only moderate task loads.

There was a significant effect from the throttle type on the perceived mental workload. The mean perceived mental workload for the automatic throttle was 49.64, while the mean for the manual throttle was 57.1, again, both are moderate task loads.

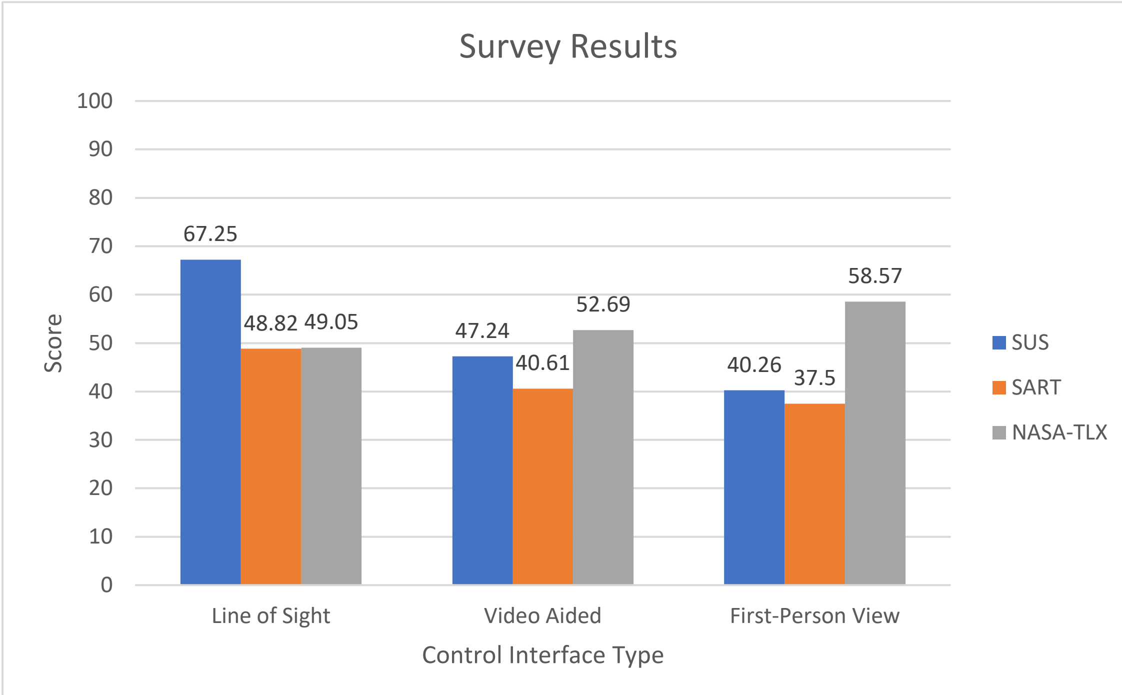

Figure 9: Summary of the survey results for each of the control interfaces.

4.3 Situational Awareness

The operator's post lap SART surveys showed a significantly higher situational awareness while using the LOS control interface (F = 15.6588, p-value = <0.0001) compared to the video aided and FPV control interfaces. As with the perceived mental workload results above a Tukey-Kramer pairwise comparison was conducted to further determine if there was any connection between any of the control interfaces. This pairwise comparison shows that there is a connection between the video aided and FPV control interfaces (meaning that they are not significantly different) while there is no connection to the LOS control interface.

Considering the number of obstacles on the course there was no significant difference in the level of situational awareness when there is two obstacles or four obstacles (F = 0.0294, p-value = 0.8640).

Table 2: Summary of results by control interface, throttle type, and course type for SART and NASA-TLX surveys.

Control Type	Throttle Type	Course Type	SA	r-TLX
LOS	**Auto**	**Simple**	41.30	52.90
		Complex	51.51	44.01
	Manual	**Simple**	48.91	51.78
		Complex	43.48	55.46
Video Aided	**Auto**	**Simple**	41.49	47.50
		Complex	44.28	51.16
	Manual	**Simple**	36.85	54.17
		Complex	40.04	57.59
FPV	**Auto**	**Simple**	40.60	55.79
		Complex	36.84	54.70
	Manual	**Simple**	36.27	61.80
		Complex	36.61	61.66

There was a significant difference in the level of situational awareness depending on the throttle type (F = 5.0734, p-value = 0.0253), and on examination of the means the use of the automatic throttle resulted in the higher reported situational awareness.

4.4 System Usability

The post-testing SUS surveys showed that the line of sight control was significantly more usable than the video aided and first-person view (F = 9.2565, p-value = 0.0003). The mean score for the line of sight control is 67.25 which, according to the adjective scale defined by Bangor et al., is a good control interface, while the mean scores for the line of sight is 47.24 which is an okay interface, and the first-person view control interface is 40.26 is a poor interface (Bangor, Kortum, & Miller, 2009).

Table 3: Summary of SUS survey results for each control interface type.

Survey	Interface Type	Mean
SUS	Line of Sight	67.25 (SD = 19.41)
	Video Aided	47.24 (SD = 20.51)
	First Person View	40.26 (SD = 21.26)

5. DISCUSSION

5.1 Discussion

The use of first-person view control resulted in a significantly longer lap time than the line of sight and video aided control interfaces. This indicates that operators were taking longer to make a decision on how to advance around the course while using first-person view control compared to the other control interfaces. The operators also collided with the environment at a marginally significantly higher rate while using first-person view compared to the line of sight and video aided control. This coupled with the SA and mental workload results showed that the operators had difficulty perceiving, understanding and projecting the future state of the position of the drone relative to the environment. This resulted in a higher perception of task demand, and thus higher mental workload. This provides insight into how operators perceive the space that they are operating in. The wider field of view associated with the line of sight control allowed for an increased understanding of where potential obstacles were on the course, while the narrowed field of view when using the video stream from the drone allowed the operator to better understand the spatial orientation of the drone. There was no significant difference between the line of sight control and the video aided control methods with respect to the flight performance, perceived mental workload, situational awareness and therefore system usability scores (table 2) can be used to better understand the operator's perceived difference in the control types. The operator's perceived mental workload was slightly higher while using the video aided control, and SA was significantly lower. This relationship indicates that the increased mental workload of using both a live video stream from the drone and the visual line of sight together reduced the operator's overall awareness of the situation. The effect of switching between the video and the visual line of sight was also apparent in the operator's system usability scores for each control method. The line of sight control was rated as a "good" interface method, while the video aided control was considered an "okay" interface method, and the first-person view interface was considered a "poor" interface. This means that the line of sight and video aided control interfaces are useful and useable as they are, but they could be improved. This further supports the conclusion that when the operator needs to switch between watching the live video stream and watching the drone, this has a negative impact on the operator mentally even though there is no significant impact on flight performance. When examining the effects of the first-person view control on flight performance, there was a significant difference for lap time when compared to line of sight and video aided control methods. There was only a marginally significant difference seen in the number of collisions when compared to the other methods. This, coupled with the significantly higher perceived mental workload and significantly lower SA while using the first-person view control suggests that operators were moving through the lap at a slower pace to better understand the environment due to the field of view being narrowed to only what can be seen from the live stream. The first-person view was also significantly less usable than the line of sight control, with a below average rating. To better inform the operator about the situation as it changes there is a need to reconcile what the operator perceives in the environment and what is actually present in the environment. One solution to this that has been proposed is the use of augmented reality, which would allow the operator to maintain line of sight, while streaming the additional information about the environment and/or situation to the operator from the drone. The application of augmented reality has been applied to flight planning and supervision, and has been found to have a positive effect on the user's perception of the position of the drone relative to objects in the environment as well as on confidence (Zollmann, Hoppe, Langlotz, & Reitmayr, 2014). This indicates that the user's mental workload could be reduced with the use of augmented reality, when evaluating flight performance and situational awareness. In addition to the control interface type, it was determined that the manner in which the operator manipulates the throttle, whether automatically or manually, significantly impacts flight performance, situational awareness, mental workload, and usability. The use of an automatic throttle in this experiment allowed the drone to launch automatically and then maintain a default altitude. The use of the manual throttle was difficult for participants to get comfortable with and because of this potentially reduced operator's confidence which is a component of perceived mental workload. As operator confidence (the performance component of the NASA-TLX survey, see Appendix 2) decreases, perceived mental workload increases. With the development of alternative control interfaces, there is a need to consider modifying and optimizing the manner in which the operators control the throttle.

Professional drone racers solely rely on the use of an FPV control interface, even though this is determined to be the least usable interface for novice operators. This is due in large part to the design of courses that are used in professional races. These courses tend to be very long winding courses, set up in stadiums and warehouses, flying at speeds up to 80 miles per hour. This makes it very difficult for an operator to maintain control of the drone using the LOS control interface, and the use of the video aided interface could possibly lead to distractions during competition when the drone is out of sight from the operator.

5.2 Limitations

Some of the possible difficulties with the study can be seen in the lack of experienced operators. The use of only novice operators could be addressed in future studies by recruiting more people and expanding to include experienced and novice operators. This would allow for more practice with the drones and a better understanding prior to experimentation.

5.3 Future Work

Future work in this area would be seen in increasing the recruitment size of the study and addressing the limitations mentioned as it pertains to drone operation experience. The study could also look at the effects of an augmented reality control interface method. Future studies could also include brain based measures, such as electroencephalography or functional near infrared spectroscopy, to better understand the underlying neural activity.

6. IMPLICATIONS

6.1 Research Implications

The results indicate that the first-person view control interface type is significantly more difficult to use compared to the traditional line of sight control that the FAA recommends, and results in significantly longer lap times compared to both the line of sight control and the video aided control interfaces. The situational awareness for this study is related to the operator's perception, understanding, and projection of the situation. The focus is to determine if the operator can perceive the obstacles in the environment and the boundaries of the course, as well as the drone's position relative to them, to also understand what is happening in real time, and then project the state of the drone and the environment in the future. This is very important when operating a consumer drone, since the practical uses a drone will often be in a changing environment with the risk of interacting with people and other aircrafts.

There is not a significant difference between the video aided control and the line of sight control with respect to flight performance and mental workload, though there is a significant difference in the situational awareness and the usability, with the line of sight control scoring higher in both areas than the video aided control. This means that if an effective means of incorporating the live video stream, and possibly other useful information such as altitude and velocity, into an easy to use augmented reality headset that would not require the operator to switch between the line of sight and the video, then it may be possible to improve all aspects of the operator's performance, though this would require further testing and the development of a suitable augmented reality interface. This study could also be applied to the development of new heads-up-displays for other autonomous vehicles, such as cars, that could provide the operator with pertinent information, while not interfering with their standard method of control, which could improve the quality of trust (intervening with automation only when it is truly necessary, otherwise allowing the automation to function) the operator has in the automation of the vehicle.

7. CONCLUSIONS

The conclusions that can be drawn from this study indicate that if the video aided control interface can be simplified to reduce the amount of switching that is required, then there would be no significant difference between video aided and line of sight control, while the first-person view control would require extensive practice, or training to master. It can be concluded that for the novice consumer drone operator, which is what many consumers are, line of sight control is the best method of control, though the video aided control is not far behind. There are no other studies that examine the effects of the three primary control interface methods on flight performance, perceived mental workload, situational awareness, and what the overall usability of the interface is. Further research on this subject is required to find the best interface method that will improve operator performance, mental workload, situational awareness, and system usability so that the risk of accidents involving drones and people, and drones and other aircrafts.

8. REFERENCES

Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 114-123.

BBC News. (2015, November 26). *Toddler's Eyeball Sliced in Half by Drone Propeller*. Retrieved from BBC News - Hereford & Worcester: http://www.bbc.com/news/uk-england-hereford-worcester-34936739

Boring, R., Ulrich, T., & Lew, R. (2016). RevealFlow: A process control visualizaiton framework. *International Conference on Augmented Cognition*, (pp. 145-156). Switzerland.

Chen, J. Y., Barnes, M. J., & Harper-Sciarini, M. (2011). Supervisory control of multiple robots: human-performance issues and user-interface design. *IEEE Transactions on Systems, Man, and Cybernetics, 41*(4), 435-454.

Cho, K., Cho, M., & Jeon, J. (2017). Fly a dron safely: evaluation of an embodied egocentric drone controller interface. *Interacting with computers*, 345-354.

Durso, F. T., & Sethumadhaven, A. (2008). Situation Awareness: Understanding Dynamic Environments. *Human Factors, 50*(3), 442-448.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 32-64.

Federal Aviation Administration. (2017, August 15). *FAA Aerospace Forecast Fiscal Years 2017-2037.* Retrieved from FAA Aerospace Forecasts: https://www.faa.gov/data_research/aviation/aerospace_forecasts/

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 904-908). Human Factors and Ergonomics Society.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Emperical and Theoretical Research. *Human Mental Workload*, 139-183.

Hooey, B. L., Kaber, D. B., Adams, J. A., Fong, T. W., & Gore, B. F. (2017). The underpinnings of workload in unmanned vehicle systems. *IEEE Transactions on Human-Machine Systems*.

John A. Taylor V. Michael P. Huerta, 15-1495 (United States Court of Appeals For The District of Columbia Circuit May 19, 2017).

Kaber, D., Jin, S., Zahabi, M., & Pankok, Jr., C. (2016). The effect of driver cognitive abilities and distractions on situation awareness and performance under hazard conditions. *Transportation Research Part F, 42*, 177-194.

LaFleur, K., Cassady, K., Doud, A., Shades, K., Rogin, E., & He, B. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface. *Journal of Neural Engineering*.

Lu, J.-L., Horng, R.-Y., & Chao, C.-J. (2013). Design and test of a situation-augmented display for an unmanned aerial vehicle monitoring task. *Perceptual & Motor Skills: Motor Skills and Ergonomics*, 145-165.

Lu, M., & Lung, C. L. (2016). Studies of AR drone on gesture control. *3rd International Conference on Materials Engineering, Manufacturing Technology and Control* , (pp. 1869-1873).

Magister, T. (2010). The Small Unmanned Aircraft Blunt Criterion Based Injury Potential Estimation. *Safety Science*, 1313-1320.

Meola, A. (2017, July 13). *Drone market shows positive outlook with strong industry growth and trends*. Retrieved from Business Insider: http://www.businessinsider.com/drone-industry-analysis-market-trends-growth-forecasts-2017-7

NASA. (2017, February 15). *NASA TLX: Task Load Index*. Retrieved from Human Systems: https://humansystems.arc.nasa.gov/groups/TLX/index.php

Parasuraman, R., Cosenzo, K. A., & de Visser, E. (2009). Adaptive automaton for human supervision of multiple uninhabited vehicles: Effects on chagne detection, situation awareness, and mental workload. *Mil. Phsychol. , 21*(2), 270-297.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 849-860.

Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review*, 61-86.

Satuf, E. N., Kaszkurewicz, E., Schiru, R., & de Campos, M. C. (2016). Situation awareness measurement of an ecological interface designed to operator support during alarm floods. *International Journal of industrial Ergonomics, 53*, 179-192.

Statista. (2017). *Sales of consumer drones to dealers in the United States from 2013 to 2017 (in million U.S. dollars)*. Retrieved from Consumer Electronics: https://www.statista.com/statistics/641932/us-consumer-drones-wholesale-sales/

Tharanathan, A., Bullemer, P., Laberge, J., Reising, D. V., & Mclain, R. (2012). Impact of Functional and Schematic Overview Displays on Console Operators' Situation Awareness. *Journal of Cognitive Engineering and Decision Making, 6*(2), 141-164.

Zollmann, S., Hoppe, C., Langlotz, T., & Reitmayr, G. (2014). FlyAR: Augmented Reality Supported Micro Aerial Vehicle Navigation. *IEEE Transactions on Visualization and Computer Graphics*, 560-568.

APPENDIX I – SITUATIONAL AWARENESS RATING TECHNIQUE SURVEY

Instability of Situation
How changeable is the situation? Is the situation highly unstable and likely to change suddenly (high) or is it very stable and straightforward (low)?

1 2 3 4 5 6 7

Complexity of Situation
How complicated is the situation? Is it complex with many interrelated components (high) or is it simple and straightforward (low)?

1 2 3 4 5 6 7

Variability of Situation
How many variables are changing within the situation? Are there a large number of factors varying (high) or are there few variables changing (low)?

1 2 3 4 5 6 7

Arousal
How aroused are you in the situation? Are you alert and ready for activity (high) or do you have a low degree of alertness (low)?

1 2 3 4 5 6 7

Concentration of Attention
How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)?

1 2 3 4 5 6 7

Division of attention

How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)?

1 2 3 4 5 6 7

Spare Mental Capacity

How much mental capacity of you have to spare in the situation? Do you have sufficient capacity to attend to many variables (high) or nothing to spare at all (low)?

1 2 3 4 5 6 7

Information Quantity

How much information have you gained about the situation? Have you received and understood a great deal of knowledge (high) or very little (low)?

1 2 3 4 5 6 7

Familiarity with Situation

How familiar are you with the situation? Do you have a great deal of relevant experience (high) or is it a new situation (low)?

1 2 3 4 5 6 7

Information Quality

How valuable is the information received form the situation? Is the quality of the information very good (high) or is the quality very poor (low)?

1 2 3 4 5 6 7

APPENDIX II – NASA-TLX SURVEY

NASA-TLX

Participant ID: ______________

Trial #: ______________

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

APPENDIX III – SYSTEM USABILITY SCALE SURVEY

System Usability Scale
Control type:

	Strongly disagree				Strongly agree
	1	2	3	4	5
1. I think that I would like to use this system frequently					
2. I found the system unnecessarily complex					
3. I thought the system was easy to use					
4. I think that I would need the support of a technical person to be able to use this system					
5. I found the various functions in this system were well integrated					
6. I thought there was too much inconsistency in this system					
7. I would imagine that most people would learn to use this system very quickly					
8. I found the system very cumbersome to use					
9. I felt very confident using the system					
10. I needed to learn a lot of things before I could get going with this system					

APPENDIX IV – DETAILED STATISTICAL ANALYSIS

Summary of Effects

Table 4: Summary of effects for each factor. The p-values of the control type and throttle type show that those two factors significantly affect the responses of lap time, number of collisions, number of course deviations, situational awareness, and perceived mental workload. The interaction of control type and throttle type shows marginally significant effects on the responses. The course type, (number of obstacles) does not significantly affect the responses, and neither do the interactions with control type and throttle type.

Source	LogWorth	PValue
Control Type	6.731	0.00000
Throttle Type	6.084	0.00000
Control Type*Throttle Type	1.212	0.06140
Course*Throttle Type	0.658	0.21984
Control Type*Course	0.560	0.27532
Course	0.369	0.42784 ^

Lap Time

Table 5: Summary of the factor effects on the lap time. The control type and the throttle type are both significant effects, and the interaction between the control type and the throttle type is marginally significant.

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Control Type	2	2	7653.140	4.6514	0.0106*
Course	1	1	24.102	0.0293	0.8643
Control Type*Course	2	2	1402.315	0.8523	0.4279
Throttle Type	1	1	19465.869	23.6618	<.0001*
Control Type*Throttle Type	2	2	4616.731	2.8059	0.0627
Course*Throttle Type	1	1	10.815	0.0131	0.9088

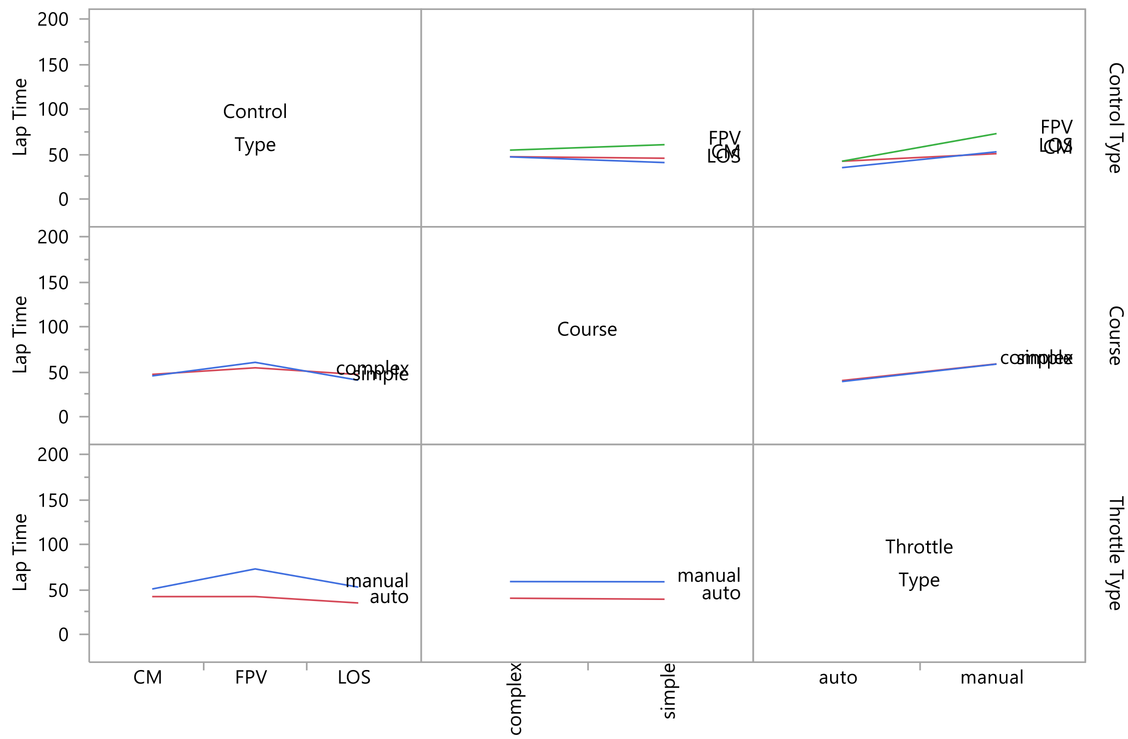

Figure 10: Factor interaction profile for the lap time. There is a slight interaction between the control type and the throttle type.

Figure 10 above is an interaction profile for the control interfaces, throttle type, and course type. This profile is used to show any interaction between the various independent variables for a specific dependent variable, in this case the lap time. If the lines for a given profile cross over then there is an interaction, while if they do not then there is no significant interaction between those independent variables.

Number of Collisions

Table 6: Summary of the factor effect on the number of collisions. The throttle type has a significant effect on the number of collisions, and the control type and the interaction between the control type and the throttle type are marginally significant.

Source	**Nparm**	**DF**	**Sum of Squares**	**F Ratio**	**Prob > F**
Control Type	2	2	32.18737	2.7415	0.0668
Course	1	1	1.75379	0.2988	0.5853
Control Type*Course	2	2	4.85660	0.4137	0.6618
Throttle Type	1	1	151.69795	25.8414	<.0001*
Control Type*Throttle Type	2	2	30.33285	2.5836	0.0779
Course*Throttle Type	1	1	8.89081	1.5145	0.2198

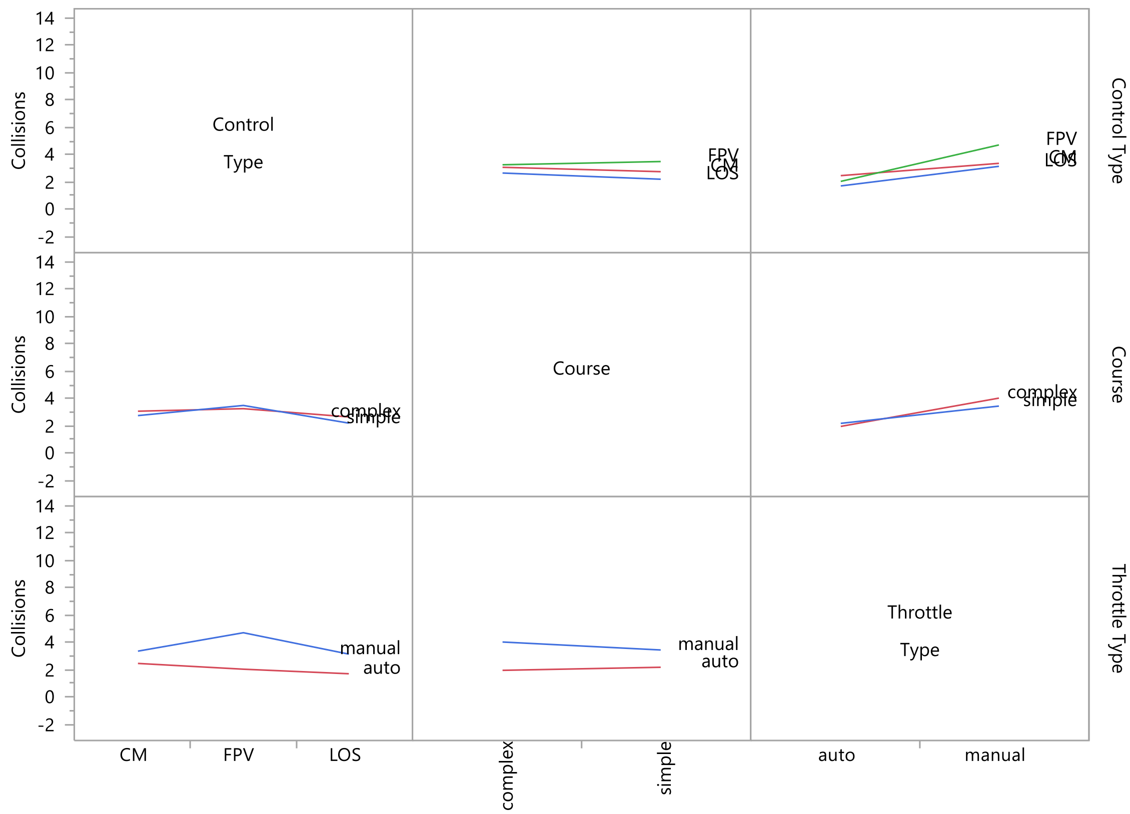

Figure 11: Factor interaction profile for the number of collisions. There is a significant interaction between the throttle type and control type, as well as a slight interaction between the course type (number of obstacles) and the throttle type.

Course Deviations

Table 7: Summary of the factor effects on the number of course deviations. There are no significant effects on the number of course deviations, but there is a marginally significant effect from the throttle type and the interaction between the throttle type and control type.

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Control Type	2	2	10.097090	1.6464	0.1952
Course	1	1	1.935403	0.6312	0.4278
Control Type*Course	2	2	3.196850	0.5213	0.5945
Throttle Type	1	1	10.353912	3.3765	0.0676
Control Type*Throttle Type	2	2	17.344659	2.8282	0.0614
Course*Throttle Type	1	1	1.082615	0.3531	0.5530

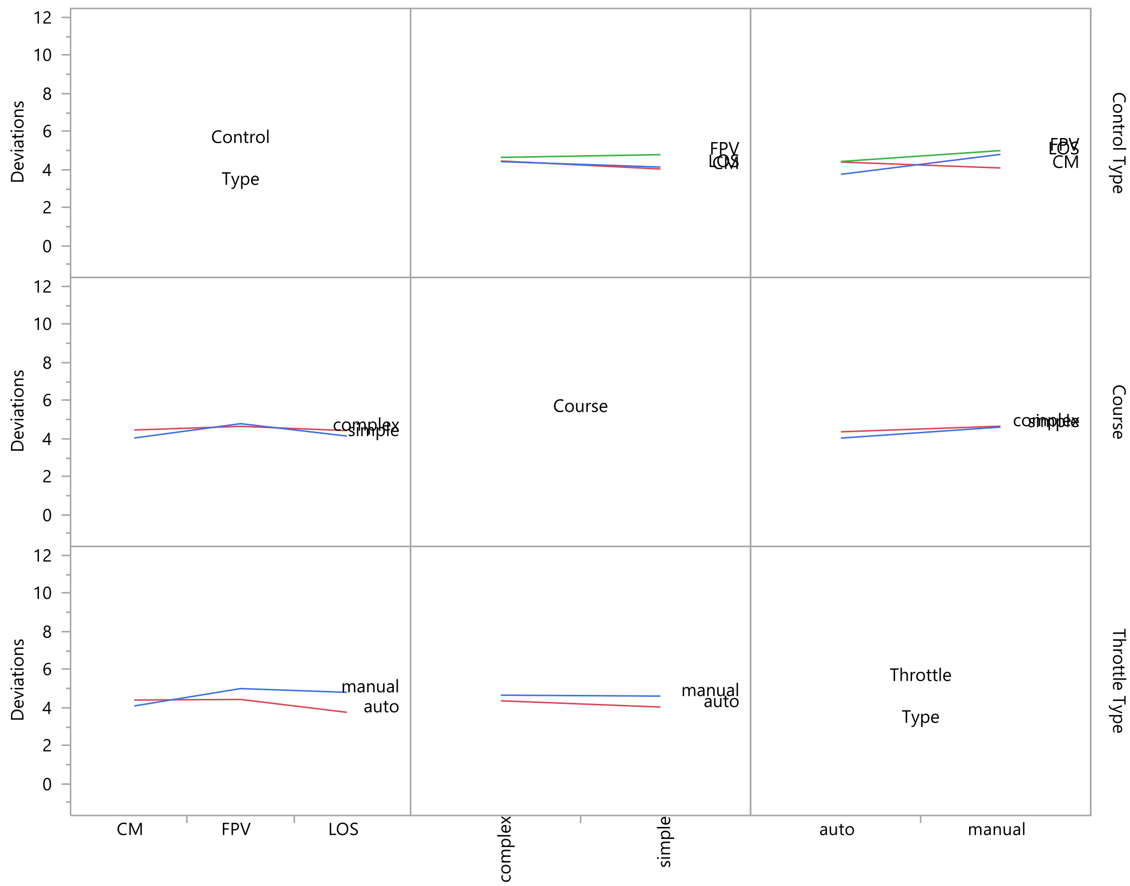

Figure 12: The factor interaction profile for the number of course deviations. There is a significant interaction between the control type and the throttle type.

Perceived Mental Workload

Table 8: Summary of factor effects on the Perceived Mental Workload. There is a significant effect from the control type and the throttle type on the perceived mental workload.

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Control Type	2	2	3343.0068	6.3857	0.0020*
Course	1	1	102.3862	0.3911	0.5324
Control Type*Course	2	2	205.3740	0.3923	0.6760
Throttle Type	1	1	2785.6534	10.6420	0.0013*
Control Type*Throttle Type	2	2	38.8114	0.0741	0.9286
Course*Throttle Type	1	1	119.7753	0.4576	0.4995

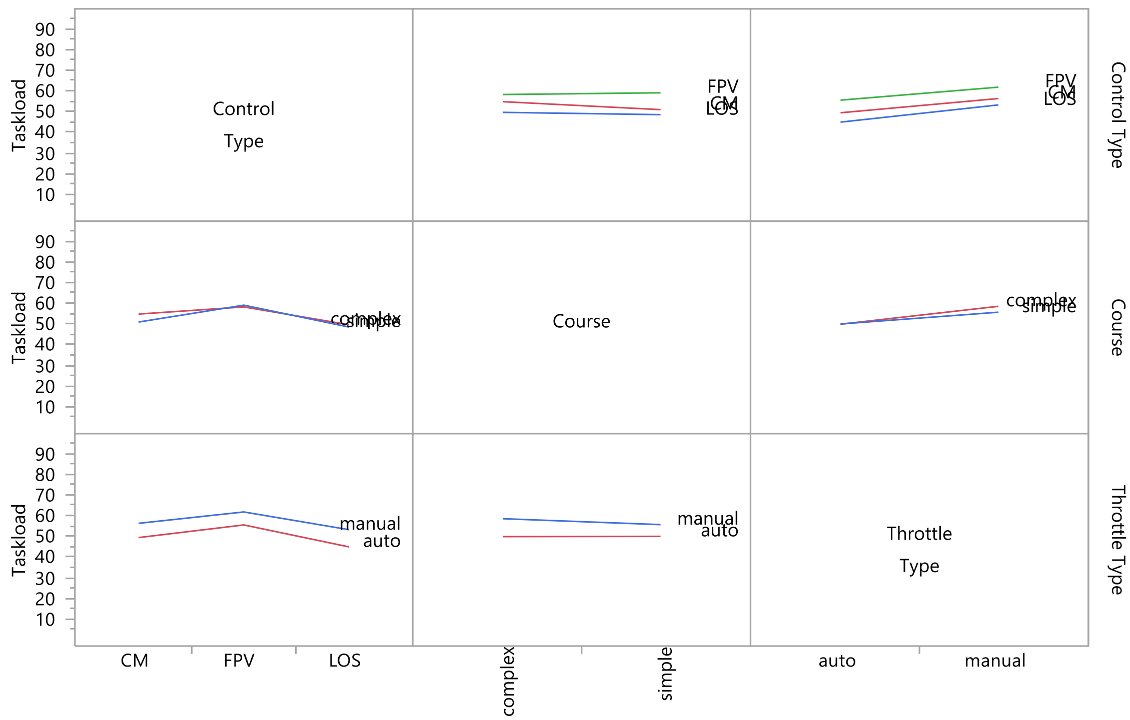

Figure 13: Factor interaction profiles of the perceived mental workload. There does not appear to be any significant interactions.

Situational Awareness

Table 9: Summary of factor effects on situational awareness. There is a significant effect from the control type and the throttle type on the situational awareness.

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Control Type	2	2	5221.4950	16.7132	<.0001*
Course	1	1	33.4086	0.2139	0.6442
Control Type*Course	2	2	405.4748	1.2979	0.2753
Throttle Type	1	1	832.3589	5.3285	0.0220*
Control Type*Throttle Type	2	2	132.1415	0.4230	0.6557
Course*Throttle Type	1	1	0.1338	0.0009	0.9767

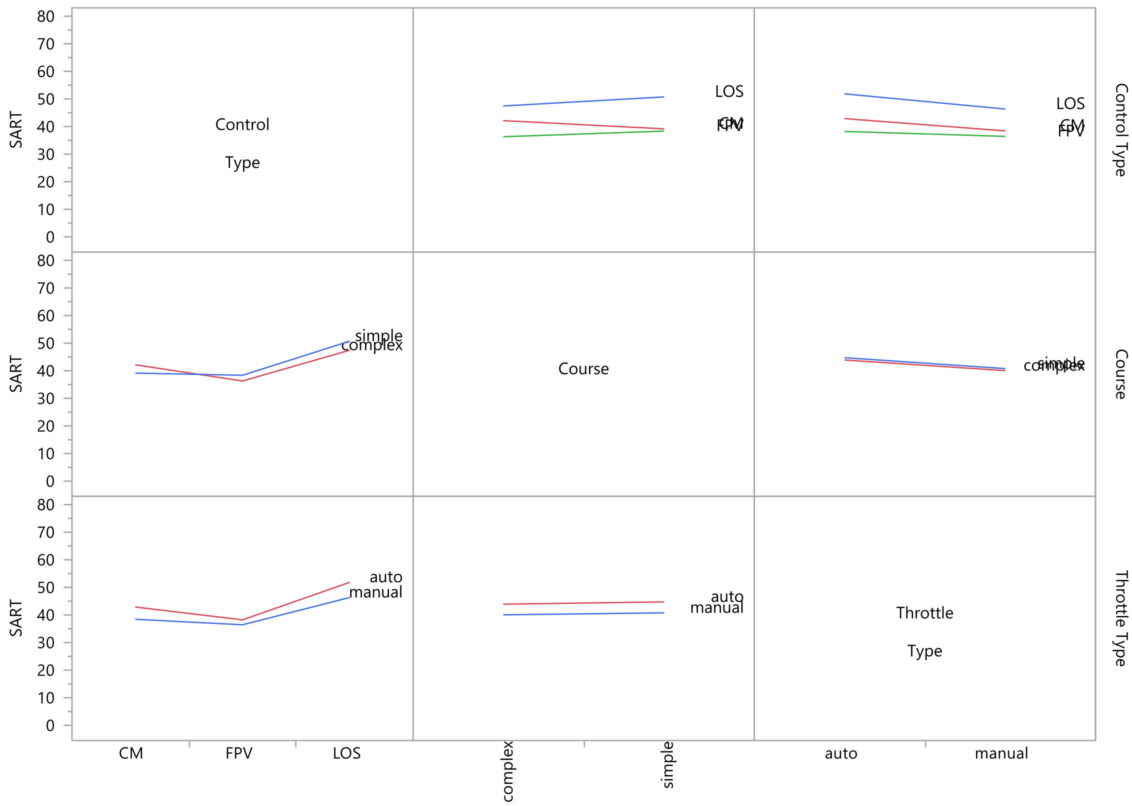

Figure 14: Factor interaction profile for the situational awareness. There does not appear to be any significant interactions.